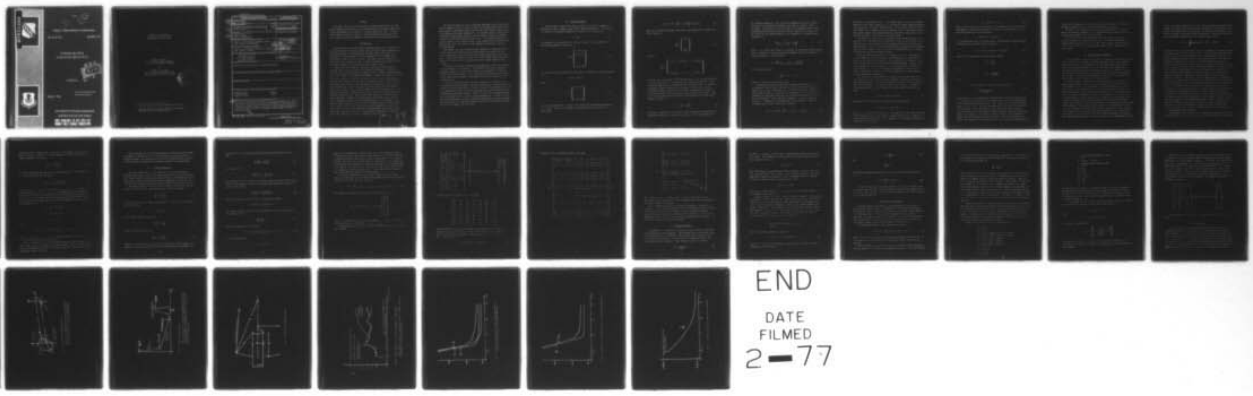
AD-A034 396 FRANK J SEILER RESEARCH LAB UNITED STATES AIR FORCE --ETC F/G 20/6
ESTIMATION AND CONTROL IN MULTIDITHER ADAPTIVE OPTICS.(U)
DEC 76 R B ASHER, R F OGRODNIK

UNCLASSIFIED FJSRL-TR-76-0015 NL

| OF |
AD
A034396



END

DATE
FILMED
2-77

FRANK J. SEILER RESEARCH LABORATORY

SRL-TR-76-0015

DECEMBER 197

ESTIMATION AND CONTROL IN MULTIDITHER ADAPTIVE OPTICS

SCIENTIFIC

PROJECT 2304

AIR FORCE SYSTEMS COMMAND

UNITED STATES AIR FORCE

ESTIMATION AND CONTROL IN MULTIDITHER ADAPTIVE OPTICS

by

[+]Robert B. Asher
F.J. Seiler Research Laboratory
U.S. Air Force Academy, CO 80840

and

[++]Robert F. Ogrodnik
Rome Air Development Center (OCTM)
Griffiss Air Force Base, NY 13441

[+] Research Associate, Applied Mathematics Division

[++] Research Scientist, Coherent Optical and Microwave High Resolution Techniques Section

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: FJSRL-TR-76-0015 AD A-
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)*: ESTIMATION AND CONTROL IN MULTIDITHER ADAPTIVE OPTICS.

5. TYPE OF REPORT & PERIOD COVERED: Scientific rept.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Robert B. Asher, Robert F. Ogrodnik

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Frank J. Seiler Research Laboratory (AFSC), USAF Academy, Colorado 80840

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: DRS61102F, 2304-F1-56, F1

11. CONTROLLING OFFICE NAME AND ADDRESS: Frank J. Seiler Research Laboratory (AFSC), USAF Academy, CO 80840

12. REPORT DATE: December 1976

13. NUMBER OF PAGES: 33

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: 35p.

15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Adaptive Optics
Estimation Theory
Kalman Filtering
Optics
Speckle

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This paper develops optimal statistical estimation formulation for multi-dither adaptive optics control loops which potentially can enhance stable beam control performance in the presence of spurious signal like noise. A multidither autofocus example is chosen to compare estimated assisted and conventional adaptive optical performance in the presence of speckle generated noise.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

319920

1B

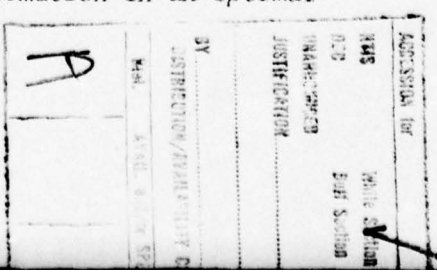
Abstract

This paper develops optimal statistical estimation formulation for multi-dither adaptive optics control loops which potentially can enhance stable beam control performance in the presence of spurious signal like noise. A multi-dither autofocus example is chosen to compare estimated assisted and conventional adaptive optical performance in the presence of speckle generated noise.

I. Introduction

Adaptive optics has truly become an important subfield. As the ideas in adaptive optics cross interdisciplinary lines of both the optics community and the control community (although perhaps not universally recognized), it is imperative that new techniques from each community evolve to enhance the performance of adaptive optical systems. This paper represents one of these inderdisciplinary crossings in order to utilize experience and techniques from the control community to enhance the field of adaptive optics.

Adaptive optical systems are inherently closed loop systems whereby a feedback signal is used to detect system error in some manner. This error detection is then used to control an adaptive optic in order to enhance overall optical performance. The inherent performance of closed loop control lies in its ability to correct for a process that may not be as accurately determined as one would like. However, the limitation of feedback control is that the feedback signals must be as precisely determined as the accuracy level required of the controller [1]. Since in adaptive optics the signal may be corrupted by several noise effects such as speckle, photon noise, and detector noise, it is imperative to obtain a filtering of this noise in order to properly achieve the potential system accuracy. Without fully realizing the use of optimal techniques, the optics community has actually benefited from a form of filtering that is inherently present in existing adaptive optics servo hardware. The inherent analog response of these servos to feedback signals yield a filtering of the signal. However, the filtering is at best suboptimal as it does not include the prior statistical knowledge of the phenomenon perturbing the feedback signal. Nor does it use any of the available information in an optimal sense. The filtering is strictly a bandpass operation.



This paper develops the use of estimation techniques as well known in the control community [2,3,4] in order to obtain an optimum estimate of the feedback signal for multidither adaptive optics (see [5,6] Coherent Optical Adaptive Techniques and Image Compensation). The feedback signal is then used in a stochastic approximation scheme in order to obtain a convergent feedback controller for the multidither adaptive optics. The estimator will give an optimal estimate of the feedback signal in the presence of both speckle noise and detector noise. The results of the technique are applied to an autofocusing scheme.

The control is based on optimizing the intensity as a function of the distance between the secondary and primary mirrors for a Cassegrain telescope. Maximization logic is based on the parameter gradients of the intensity. Second order gradients are not used in this simple maximization logic. However, they can easily be incorporated in multi-element adaptive optic systems.

The worse case speckle problem has been addressed. This is the additive speckle problem. In the example, a simple phase lock loop was modeled for comparison purposes. An AGC compensation was not simulated as it does not add any utility to the models. It is shown that the technique in this paper yields significant performance improvement even with ideal modeling of the comparison loop.

Section I gives the problem statement including a discussion of the sinusoidal perturbation adaptive controller and the basic estimation equations. Section II considers the problem of speckle interaction on adaptive optics from the standpoint of induced signal spectrum components, and the modeling of an estimator addressing this problem. Section III gives the estimator equations in the continuous form for possible analog implementation. Section IV gives the control philosophy in terms of stochastic approximation. Section V demonstrates the results of applying estimation theory to the autofocusing problem and Section VI contains the conclusion.

II. Problem Statement

The philosophy of multidither adaptive optics is to obtain a maximum of intensity by using a feedback signal derived from multidithering. The intensity is a function of adjustable parameters, P, where P is a q-vector, i.e.,

$$I = I(P). \tag{1}$$

The parameters are perturbed from a reference value, $\bar{P}$, by the addition of a sinusoidal perturbation signal, i.e.,

$$P = \bar{P} + \begin{pmatrix} c_1 \sin \omega_1 t \\ c_2 \sin \omega_2 t \\ \cdot \\ \cdot \\ \cdot \\ c_q \sin \omega_q t \end{pmatrix}. \tag{2}$$

The actual intensity at the receiver will then be a function of the perturbations, i.e.,

$$I(P) = I(\bar{P} + \Omega) \tag{3}$$

where

$$\Omega = \begin{pmatrix} c_1 \sin \omega_1 t \\ \cdot \\ \cdot \\ \cdot \\ c_q \sin \omega_q t \end{pmatrix}. \tag{4}$$

Eq. (3) may be expanded in a Taylor series about small amplitude perturbations ($||\Omega|| << ||\bar{P}||$ where $||(\cdot)||$ denotes the Euclidean norm of the vector $(\cdot)$). This yields

$$I(P) = I(\bar{P}) + \frac{\partial I^T}{\partial P}\,\Omega + \frac{1}{2}\,\Omega^T\,\frac{\partial^2 I}{\partial P^2}\,\Omega + \text{H.O.T.} \tag{5}$$

where H.O.T. denotes the higher order terms, the superscript "T" denotes the vector transpose,

$$\frac{\partial I}{\partial P} = \begin{pmatrix} \partial I/\partial P_1 \\ \cdot \\ \cdot \\ \cdot \\ \partial I/\partial P_q \end{pmatrix} \tag{6}$$

and where

$$\frac{\partial^2 I}{\partial P^2} = \begin{pmatrix} \partial^2 I/\partial P^2{}_1,\ \partial^2 I/\partial P_2 \partial P_1,\ \ldots,\ \partial^2 I/\partial P_q \partial P_1 \\ \cdot \\ \cdot \\ \cdot \\ \partial^2 I/\partial P_1 \partial P_q,\ \ldots\ldots\ldots\ldots,\ \partial^2 I/\partial P^2{}_q \end{pmatrix}. \tag{7}$$

P_i denotes the i-th element of the parameter vector. For small perturbations (in terms of the norm of Ω) the higher order terms may be dropped. A more refined approach will be considered when the second derivative is employed for convergence behavior monitoring in this paper. For the purpose of this development, however, analog filtering will be assumed to be available for eliminating the double frequency terms as well as higher order terms. Therefore, if the signal return were perfect, we would have as the signal after bandpass filtering

$$I(P) = I(\bar{P}) + \frac{\partial I^T}{\partial P}\,\Omega. \tag{8}$$

The sign and magnitude of the gradient of the intensity with respect to the adjustable parameters are contained explicitly in the term containing the

basic dither frequencies. This yields the information necessary to make a correction to the adjustable parameters such that the intensity can be increased. In particular, the new settings for the parameters may be found by appropriately choosing the desired intensity change at the current time and solving for the required change in the parameters by a steepest ascent algorithm, i.e.,

$$\bar{P}_{new_{k+1}} = \bar{P}_{old_k} + \varepsilon_k \frac{\partial I}{\partial \bar{P}_k} \tag{9}$$

where $\varepsilon_k > 0$ at each time instant, k, and serves as an effective gain parameter that is controlled to assure reasonable convergence towards the optimum. The predictor change in I may be found as

$$\delta I = \frac{\partial I^T}{\partial \bar{P}} (\bar{P}_{new} - \bar{P}_{old}) = \varepsilon_k \frac{\partial I^T}{\partial \bar{P}} \frac{\partial I}{\partial \bar{P}} . \tag{10}$$

This is continued until

$$|| \frac{\partial I}{\partial \bar{P}} || < \delta, \quad \delta > 0 \tag{11}$$

where δ is chosen for convergence control.

The above algorithm implicitly used in multidither Adaptive Optics has the problem of the gradient so obtained is assumed to be a deterministic quantity and as such is determined by a noiseless measurement. Neither of these cases are actually realized in any application. In particular, in real situations the amplitude of the gradient is amplitude modulated by speckle as well as corrupted by signal like noise components, as will be showed in Section II. Furthermore, the measurements are not noiseless. Thus, the returned signal from the target may be written as

$$y(I) = k(z)\{I(\bar{P}) + \frac{\partial I^T}{\partial \bar{P}} \Omega + s_1 + s_2 + \ldots + s_q + T + g\} + n \tag{12}$$

where $k(z)$ is the detector gain, z is the target range, $\partial I/\partial P$ is the gradient modulated at the dither frequency, s_i is the speckle induced spectrum component occurring at the i-th dither frequency band, T is the target diffusely distributed return, g is a random glint return component and η is the detector noise. The speckle induced spectrum and modulation at the dither frequencies cause a random fluctuation of the signal at each dither band.

The second order terms may be retained with some further modeling. The intensity $I(\bar{P})$ is assumed to be representative of the returned intensity with g and T being zero-mean processes for the temporal variation of the glint and diffuse returns. In this way, the temporal spectral spreading of the returned intensity may be modeled. That is, g and T are perturbations of the returned intensity due to glint and diffuse changing.

Thus, the measurement equation is as given in Eq. (12). Since speckle induced components which effect system performance appear at the dither frequencies, they are indistinguishable from the gradient vector components and cannot be separated out by conventional filtering techniques. However, by the use of optimal estimation techniques, it can be discriminated as undesirable signal components. In fact, an estimator is structured in Section III in order to estimate the gradient, the speckle components, the glint component, and the target diffuse component.

The structure of the estimator to be used is given without proof as it is well known to the control and estimation community. However, the structure is given for completeness. The first formulation is that of an analog implementation of the estimator. The process x, which is an n-vector, is assumed to evolve according to the set of formal first order vector differential equations,

$$\dot{x}(t) = f\{x(t)\} + G(t)u(t) \tag{13}$$

where $u(t)$ is a zero mean white noise m-vector with covariance

$$E\{u(t)u(\tau)^T\} = Q(t)\delta(t-\tau)$$

where $\delta(\cdot)$ is the Dirac delta function. (The authors are well aware of the Ito interpretations as well as the general structure of the result. However, it is formalized for the sake of the optics community.) The measurement equation for the process $x(t)$ is given as

$$y(t) = h(t)^T x(t) + v(t) \tag{14}$$

where y may be a vector in general but will be assumed to be a scalar since the measurement of the process of concern is a scalar. The vector v(t) is a zero mean white noise process with covariance

$$E\{v(t)v(\tau)\} = r(t)\delta(t-\tau) \ .$$

It is assumed that u and v are uncorrelated. The approximate conditional mean estimator for the nonlinear system is given as

$$\dot{\hat{x}}(t) = F(t)\hat{x} + K(t)\{y(t) - h^T(t)\hat{x}(t)\} \tag{15}$$

where $\hat{x}$ is the approximate conditional mean estimate,

$$F(t) = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}} , \tag{16}$$

and

$$K(t) = \frac{P(t)h(t)}{r(t)} \tag{17}$$

with

$$\dot{P}(t) = F(t)P(t) + P(t)F(t)^T + G(t)Q(t)G(t)^T - \frac{P(t)h^T(t)h(t)P(t)}{r(t)} \tag{18}$$

The matrix P(t) is the approximate covariance matrix for the estimation error, $x - \hat{x}$, and the last term in Eq. (18) represents the effect of the measurements on the differential equation for P(t). The structure is called the extended Kalman estimator [2,3]. Note that the matrices F, G, Q, and r do not have to be time varying. It is important to realize that the estimator weighs the process dynamics and the measurements according to their statistical knowledge. Also, as the number of measurements processed increases, if the estimator is stable, then the knowledge gained will increase. This is significant since some of the signal processing algorithms in use in currently configured optical

systems lose information by not taking this fact into account. The discrete estimation algorithm may be easily established and is not given.

The above is an elementary discussion of estimation theory. For more details and advanced techniques, it is suggested that the reader use the previously cited references for more information.

It suffices to say that in order to obtain an estimator that will optimally estimate the quantities in Eq. (12), it is necessary to obtain a dynamic model structure as in Eqs. (13-18). Thus, dynamic statistical equations for each term in Eq. (12) must be obtained. This is considered in the next section.

III. Speckle Effects Modeling

As was indicated in the last section, it is necessary to take into account all the statistical information available about the process to be estimated. Part of the statistical information is derived from a priori models for the time varying random phenomenon. This a priori statistical information is contained in the models of the form as given in Eq. (14). Thus, in order to use the structure appearing in the optimal estimator, we must develop models such as these. This will be accomplished in this section for speckle interactive effects on the adaptive optics case in question.

We are interested in obtaining a general expression for the temporal power spectral density (PSD) for frequency components generated by target return dynamic speckle. For the purposes of this development, target backscatter can simply be characterized as having a specular and a diffuse part. The specular return preserves the spatial coherence of the incident illumination. It eminates from target regions which are normal to the receiver line-of-sight and have a smooth surface texture relative to the dimensions of the incident wavelength. The diffuse part eminate from spatially distributed target regions within the beam illumination and are characterized by the general surface roughness properties and target geometry. Both Goldfischer [7] and Crane [8] develop relationships which give the spatial PSD, $S(\omega,\Omega)$, in terms of the convolution of the object illumination function $p(u,v)$. We will use these relations, addressing the projected illumination function (of an otherwise three dimensional target shape, and resulting target illuminated structure)

in the (u,v) plane, denoted p(u,v), as the source of both specular and diffuse returns. The object shape dictated normal incident point is located in Fig. 1 at the (u,v) plane origin for convenience. The remaining distributed regions in p(u,v) account for the object diffuse backscatter. Employing Goldfischer's relatively simple formulation for the Fraunhofer case, we have

$$S(\omega,\Omega) \approx \iint_{-\infty}^{\infty} du\,dv\,p(u,v)\,p\left(u - \frac{\lambda h\omega}{2\pi},\ v - \frac{\lambda h\Omega}{2\pi}\right). \tag{19}$$

Or simply stated, the spatial PSD is the self convolution of the target spatial irradiance (refer to [7] and [8] for proof). The spatially distributed amplitude of p(u,v) is directly relatable to the target specular regions and the overall backscatter diffuse nature of the reflecting surface. The latter is generally characterized by the fall off of backscattered intensity as a function of incidence angle. The results are a combined effect of the target surface scattering properties and target illuminated shape. Surface scattering can be measured in the laboratory as Mono-Directional or Bi-Directional Reflectance data (MDR or BDR) for flatplate samples of the target surface material, giving scattering behavior which is free of shape effects (no geometrical effects). This point is made to emphasize that the formulation of p(u,v) must contain the full phenomenon of surface and shape scattering for any particular object given a specified illumination function on that object. This is a formidable task from an analytical approach for complex target structures, but reasonably performed via computer target shape decomposition techniques where laboratory BDR data can be used directly for scattering calculations. The following discussion of shape responsive spatial PSD uses this object shape decomposition approach. This involves separating a complex structured target into its fundamental geometrical shapes, identifying the projected silhouette of each shape onto the (u,v) plane, weighting this projection by the beam illumination outline, and noting normal incident point (referenced as the source of target scattered glints) locations on each beam weighted shape silhouette.

The mechanics of the computer shape decomposition approach for spatial PSD calculations, simply utilize the above stated relationship of p(u,v) and

$S(\omega,\Omega)$ by performing the self convolution and the cross correlation of each beam weighted shape silhouette with itself and with all other illuminated shapes. An important consideration which must be given in the case of adaptive optics is the effective differing target illumination functions which occur in adaptive beam forming control versus the total illumination which is sensitive to the adaptive convergence condition at any instant of time. For the multi-dither example considered here with separate (segmented) adaptive subapertures (a multidither zonal implementation caseB, two illuminating functions must be considered. These are the target illumination patterns associated with each subaperture dithered channel, driven at some dither frequency ω_i for i = 1, ... q dither channels, and the coherent adaptive control summed illumination which results, at the target and is sensitive to the adaptively controlled instantaneous convergence state.

Considering the coherently summed illumination pattern first, the resulting illumination weighted target shape silhouette are checked for normal incidence points (target glints and specular sources). The self convolution of a glint (having a (u,v) plane spatially infinitesimal value) yields a highly peaked but very low spatial frequency band contribution to the PSD, as shown in Fig. 2. The self convolution of shape silhouettes (being spatially extended in the (u,v) plane) is lower in magnitude but more extensively spread in the frequency domain as shown. The cut off of this spread is inversely related to the longest dimension of the (u,v) plane projection of the target silhouette. The cross correlation of the dither subaperture illumination with the resultant summed illumination is in general spatially broader than the self convolved summed illumination but much lower in peak amplitude due to the multidither implementation approaches which modulates a very small fraction of the total beam power at each dither frequency. This modulation is typically less than 10%, and is usually in the order of 2 to 3%. With reference to the log scale employed in Fig. 2, the cross-correlation terms will be typically 15 to 17 db below the glint contributing terms.

As stated earlier, what is needed is the temporal PSD. Since speckle motion is by relative target-beam motion, the spatial PSD will be temporally modulated by any target or beam dynamics. This can be as complicated as beam apparent motion occurring during adaptive control (adaptive search and lock-up dynamics for a glint point or a designated target location), or a

simply dictated by target spin or apparent angular rate due to flight dynamics. If motion effects are simply associated with an effective target rotation, $\dot{\theta}$, then the temporal PSD can be simply derived from the spatial PSD by multiplying the self-convolution components by $2\dot{\theta}h$. The cross-correlation terms are modulated by the dither carrier frequencies, f_i, and the effective $\dot{\theta}$ which generates a frequency spread in the temporal PSD components related to the multidither subaperature contributions. The resulting temporal PSD is what is displayed in Fig. 2. Only a point receiver case is chosen. For any specific system configuration, the PSD must be further weighted by the spatial transfer function associated with the receiver diameter. This will result in a reduced spectrum content in the speckle induced PSD and modification of certain frequency component peaks. However, in order to address the worse case condition and to choose a general formulation, a point receiver is more appropriately modeled here. In this modeling case, the important terms are the various spectral peaks associated with the glint, diffuse, and cross correlated dither signal like noise contributing terms. These are shown in Fig. 2.

Thus, Fig. 2 illustrates the temporal power spectrum found. The temporal spectrum may now be used to find the necessary state models as in Eq. (13). The procedure is to find an analytical and rational expression for the temporal PSD and then to use the well known expression

$$S_{yy}(\omega) = |H(j\omega)|^2 S_{wn}(\omega) \tag{20}$$

relating the required PSD (S_{yy}) to a linear shaping filter ($H(j\omega)$) driven by unity variance white noise. The realization of $H(j\omega)$ will give us the required differential equations.

The PSD for each of the speckle processes are now presented. For the glint, it is assumed that the PSD is

$$S_{gg}(\omega) = \frac{2\beta_g \sigma_g^2}{\omega^2 + \beta_g} \tag{21}$$

where $2\sigma_g^2/\beta_g$ is the zero frequency amplutide of the glint and β_g is the correlation constant (equal to the inverse of the correlation time). The

correlation time is assumed large. This gives a low frequency cut off for the model as the physics depicts. A similar model may be deduced for the diffuse induced component. This yields

$$S_{TT}(\omega) = \frac{2\beta_T \sigma_T^2}{\omega^2 + \beta_T^2} . \tag{22}$$

The dither correlated components require a different model. These models are of the form for each dither frequency

$$S_{s_i s_i}(\omega) = \frac{\sigma_i^2}{(\omega^2 - \omega_i^2) + \beta_i^2 \omega^2} \tag{23}$$

where β_i, i = 1, ..., q are of the order of β_g, ω_i, i = 1, ..., q are the dither frequencies, and σ_i^2, i = 1, ..., q are amplitude scaling parameters. These may be obtained as an approximation to the actual PSD curves found from Figure 2. They are adequate rational approximants to the actual PSDs. Use of Eq. (20) and spectral factorization yields the following differential equations for the models, i.e.,

$$\begin{aligned} \dot{g} &= -\beta_g g + \sqrt{2\beta_g}\, \sigma_g u_g , \\ \dot{T} &= -\beta_T T + \sqrt{2\beta_T}\, \sigma_T u_T , \\ \dot{S}_i &= x_{2_i} , \end{aligned} \tag{24}$$

and

$$\dot{x}_{2_i} = -\beta_i x_{2_i} - \omega_i^2 S_i + \sigma_i u_i, \quad i = 1, \ldots, q$$

where u_g, u_t, and u_i, i = 1, ..., q are zero mean white noise with unity variance and uncorrelated with each other.

The coefficients may be determined through physical knowledge, experimental data, or by on-line use of adaptive estimation theory. The use of this technique will yield the optimal estimator. The coefficients may be learned as the adaptive optics system is utilized.

Thus, equations (21-24) yield the temporally varying statistical information about the speckle process in a form suitable for inclusion into the optimal estimator. The next section discusses the remaining models necessary for the estimator and develops the estimator equations for implementation.

III. Estimation Equations

The basic problem that is to be addressed is that of estimating the gradient sector $\partial I/\partial P$ as this is the signal necessary to adjust the parameters. This section develops the remaining models necessary to accomplish this estimation as well as the estimation of the speckle components, and develops both the continuous estimator and the discrete estimator for this class of problems.

The remaining models necessary for the estimator structure are temporal models for $I(\bar{P})$ and $\partial I/\partial P$. From equation (10) it may be noted that

$$\frac{dI}{dt} = \frac{\partial I}{\partial \bar{P}}^T \frac{dP}{dt} \tag{25}$$

which gives the necessary dynamic equation for I. If in Eq. (9) the parameter may be written as

$$\varepsilon_k = \bar{\varepsilon}_k \Delta t$$

then a limiting process yields that

$$\frac{dP}{dt} = \bar{\varepsilon}_k \frac{\partial I}{\partial P_k} \tag{26}$$

and Eq. (32) may be written as

$$\frac{dI}{dt} = \bar{\varepsilon}_k \frac{\partial I}{\partial \bar{P}}^T \frac{\partial I}{\partial \bar{P}} . \tag{27}$$

There are several methods that may be used to model the dynamic change in the gradient. It may be noted that since the second partials are modulated into

the signal (Eq. (5)), appropriate processing will yield the matrix $\partial^2 I/\partial P^2$. Then

$$\frac{d}{dt}\left(\frac{\partial I}{\partial P}\right) = \frac{\partial^2 I}{\partial P^2}\frac{dP}{dt} \tag{28}$$

or, using Eq. (33)

$$\frac{d}{dt}\left(\frac{\partial I}{\partial P}\right) = \bar{\varepsilon}_k \frac{\partial^2 I}{\partial P^2}\left(\frac{\partial I}{\partial P}\right) \tag{29}$$

Since the matrix $\partial^2 I/\partial P^2$ cannot be extracted noiselessly, process noise must be added to this equation in order to assure the proper uncertainty level in this equation. This yields

$$\frac{d}{dt}\left(\frac{\partial I}{\partial P}\right) = \bar{\varepsilon}_k \frac{\partial^2 I}{\partial P^2}\left(\frac{\partial I}{\partial P}\right) + w \tag{30}$$

where w is assumed to be white noise, zero mean with variance

$$E\{w(t)w^T(\tau)\} = Q_w(t)\delta(t-\tau) \ .$$

Now, another dynamic model for short term modeling for the dynamic change in the gradient may be

$$\frac{d}{dt}\left(\frac{\partial I}{\partial P}\right) = - v \tag{31}$$

where each element of v is a process

$$\dot{v}_i = -\beta_{v_i} v_i + \sqrt{2\beta_{v_i}}\,\sigma_i u_{v_i}, \quad i = 1, \ldots, g \tag{32}$$

with initial condition

$$v_i = v_{o_i} > 0$$

where u_v is zero mean unity variance white noise. The correlation time is chosen on the order of the expected convergence time. The initial condition is chosen to represent the negative change in the gradient for a convergent process. The choice of correlation time as suggested will yield a model that gives a decay of the gradient from some positive value to a fluctuation about zero caused by the process noise. This model for short time intervals allows for a dynamically varying gradient vector. This effect will be studied in more detail when the estimator has been developed.

Eqs. (28-32) yield the necessary dynamic models for the estimator. It may be noted that this system of differential equations may be placed into the form of Eq. (13) where the state vector x is

$$x^T = \{I(\bar{P}),\ \nabla_p I^T,\ s_1,\ \ldots s_q,\ x_{2_1},\ \ldots x_{2_q},\ v, T, g\}^T \ . \tag{33}$$

The measurement equation may be written in the form of Eq. (12) as

$$y(t) = k(z)\{1, \Omega^T, 1, 1, \ldots, 1, 0, 0, 1, 1\} \begin{bmatrix} I(\bar{P}) \\ \nabla_p I \\ s \\ x_2 \\ v \\ T \\ g \end{bmatrix} + n \tag{34}$$

where s is the vector consisting of the elements $s_1, \ldots, s_q$ and x_2 is the vector consisting of the elements $x_{2_1}, x_{2_2}, \ldots, x_{2_q}$.

The structure as given in Eqs. 24-32 may be applied to Eq. (15) to yield the estimator

$$\begin{bmatrix} \dot{\hat{I}}(\bar{P}) \\ \nabla_p \dot{\hat{I}} \\ \dot{\hat{s}} \\ \dot{\hat{x}}_{2_1} \\ \vdots \\ \dot{\hat{x}}_{2_q} \\ \dot{\hat{v}} \\ \dot{\hat{T}} \\ \dot{\hat{g}} \end{bmatrix} = \begin{bmatrix} \varepsilon_k \hat{\nabla}_p I^T \hat{\nabla}_p I \\ -\hat{v} \\ \hat{x}_2 \\ -\beta_1 \hat{x}_{2_1} - \omega_1^2 \hat{s}_1 \\ -\beta_2 \hat{x}_{2_2} - \omega_2^2 \hat{s}_2 \\ \vdots \\ -\beta_q \hat{x}_{2_q} - \omega_q^2 \hat{s}_q \\ -\beta_v \hat{v} \\ -\beta_T \hat{T} \\ -\beta_g g \end{bmatrix} + K\left\{ y - k(z)[1, \Omega^T, 1, 1, \ldots, 1, 0, 0, 1, 1] \begin{bmatrix} \hat{I}(\bar{P}) \\ \hat{\nabla}_p I \\ \hat{s} \\ \hat{x}_2 \\ \hat{v} \\ \hat{T} \\ \hat{g} \end{bmatrix} \right\} \tag{35}$$

where the gain K is computed via Eq. (17) where

$$P = \begin{bmatrix} P_I & P_{12} & P_{13} & P_{14} & P_{15} & P_{16} & P_{17} \\ P_{12} & P_{\nabla I} & P_{23} & P_{24} & P_{25} & P_{26} & P_{27} \\ P_{13} & P_{23} & P_s & P_{34} & P_{35} & P_{36} & P_{37} \\ P_{14} & P_{24} & P_{34} & P_{x_2} & P_{45} & P_{46} & P_{47} \\ P_{15} & P_{25} & P_{35} & P_{45} & P_v & P_{56} & P_{57} \\ P_{16} & P_{26} & P_{36} & P_{46} & P_{56} & P_T & P_{67} \\ P_{17} & P_{27} & P_{37} & P_{47} & P_{57} & P_{67} & P_g \end{bmatrix} \tag{36}$$

with the element P_I being the covariance of the estimation error for I, the diagonals $P_{\nabla I}$ being the covariance of the estimation error for ∇I, and so forth. The cross terms represent correlations between the variables. The row vector h in Eq. (17) is

$$h = k(z)[1, \Omega^T, 1, 1, \ldots, 1, 0, 0, 1, 1] \; . \tag{37}$$

The matrix P may be calculated from Eq. (18) where

$$\frac{\partial f}{\partial x} = \begin{bmatrix}
0, 2\bar{\varepsilon}_k \hat{\nabla_1 I}, \ldots, 2\bar{\varepsilon}_k \hat{\nabla_q I}, & 0 \ldots 0, & 0 \ldots 0,0,0,\ldots 0, & 0, & 0 \\
0, \ldots 0, & 0 \ldots 0, & 0 \ldots 0,-1,0,\ldots 0, & 0, & 0 \\
0, \ldots 0, & 0 \ldots 0, & 0 \ldots 0,0,-1,\ldots 0, & 0, & 0 \\
 & \vdots & & & \\
0, \ldots 0, & 0 \ldots 0, & 0 \ldots 0,0,0,\ldots -1, & 0, & 0 \\
0, \ldots 0, & 0 \ldots 0, 1, & 0 \ldots 0,0,0,\ldots 0, & 0, & 0 \\
0, \ldots 0, & 0, 1, & 0 \ldots 0,0,0,\ldots 0, & 0, & 0 \\
 & \vdots & & & \\
0, \ldots 0, & 1, 0 \ldots 0, & 0 \ldots 0,1,0,\ldots 0, & 0, & 0 \\
0, \ldots 0, & -\omega_1^2, 0 \ldots 0, & -\beta_1, 0 \ldots 0, & 0 \ldots 0, & 0, 0 \\
 & \vdots & & & \\
0, \ldots 0, & -\omega_q^2, 0 \ldots 0, & -\beta_q, 0 \ldots 0, & 0, & 0 \\
 & \vdots & & & \\
0, \ldots 0, & 0 \ldots 0, & -\beta_{V_1}, & 0 \ldots 0, & 0, 0 \\
 & \vdots & & & \\
0, \ldots 0, & 0 \ldots 0, & 0, \ldots -\beta_{V_q}, & 0, & 0 \\
0, \ldots 0, & 0 \ldots 0, & 0 \ldots 0, -\beta_T, & 0 & \\
0, \ldots 0, & 0 \ldots 0, & 0 \ldots 0, & -\beta_g &
\end{bmatrix} \tag{38}$$

$$G(t) = \begin{bmatrix} 0, 0, \ldots 0, 0, \ldots 0, 0, 0 \\ \vdots \\ 0, 0, \ldots 0, 0, \ldots 0, 0, 0 \\ \sigma_1, 0, \ldots 0, 0, \ldots 0, 0, 0 \\ \vdots \\ 0, 0, \ldots \sigma_g, 0, \ldots 0, 0, 0 \\ 0, 0, \ldots 0, \sqrt{2\beta_{v_1}}\,\sigma_{v_1}, 0, \ldots 0, 0, 0 \\ \vdots \\ 0, 0, \ldots 0, 0, \ldots \sqrt{2\beta_{v_q}}\,\sigma_{v_q}, 0, 0 \\ 0, 0, \ldots 0, 0, \ldots 0, \sqrt{2\beta_T}\,\sigma_T, 0 \\ 0, 0, \ldots 0, 0, \ldots 0, 0, \sqrt{2\beta_g}\,\sigma_g \end{bmatrix} \tag{39}$$

and Q equal to a 2q + 2 identity matrix. It may be noted that $\partial f/\partial x$ is a sparse matrix. This fact may be used to eliminate many of the interconnections and, thus, simplify the equations.

This structure will yield the estimate of the physical variables to be estimated. Higher order approximations to the conditional mean may be obtained as well as the partial-integral equation of the conditional probability density function. However, the complexity of the structure is not worth the additional effort. Thus, the above yields the extended Kalman estimator. The next section considers a control philosophy.

IV. Control Philosophy

The problem is to maximize the intensity at the far field by adjustment of the pertinent system parameters. These parameters may include phase biases in segmented mirrors or a deformable mirror and the focal length of a telescope among other parameters. The intensity is a stochastic process due to perturbations of this phenomena. The performance index J is given as

$$J(P) = \max_{P \leq \mathcal{P}} E\{I(p)\} \tag{40}$$

where $\mathcal{P}$ is a constraint set for P and $E\{\cdot\}$ denotes the expected value over all realizations. The condition for optimization is that the stochastic gradient must be equal to zero, i.e.,

$$E\{\nabla_p I\} = 0 \tag{41}$$

where interchange of expectation and differentiation has occurred. The estimation algorithm yields a best estimate of this stochastic gradient. Thus, the algorithm to adjust the parameters in order to accomplish the hill climbing to reach the stationary point as depicted in Eq. (9) is

$$P_{k+1} = P_k + \varepsilon_k \hat{\nabla_p} I \tag{42}$$

where ε_k is a given sequence $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, \varepsilon_{k+1}, \ldots\}$ to be chosen and $\hat{\nabla_p} I$ is the estimator estimate for $\nabla_p I$.

The gain sequence, ε_k, must be chosen to assume convergence of the algorithm. Furthermore, it is desirable that the convergence by accelerated. The methodology for choosing the gain sequence comes from stochastic approximation [9]. (This is not an exhaustive bibliography. See [10] for a more complete bibliography.) This paper is not intended to survey all existing stochastic approximation methods. However, it is sufficient to point out that the gain sequence need not be exceedingly complex. For example, if the stochastic gradient $E\{I(P)|P\}$, where it is conditioned on P, is such that

$$\inf_{\varepsilon < ||P-P^o|| < \varepsilon^{-1}} [(P-P^o)^T E\{I(P)|P\}] > 0, \quad \forall \varepsilon > 0 \tag{43}$$

and if the estimated gradient is such that

$$E\{||\hat{\nabla_p} I||^2\} \leq h(1+||P-P^o||^2), \quad h > 0 \tag{44}$$

where P^o is the value of the parameters maximizing the intensity, then the sequence of gains that satisfy

$$\varepsilon_k \geq 0, \sum_k \varepsilon_k = \infty \tag{45}$$

and

$$\sum_k \varepsilon_k^2 < \infty$$

will yield an algorithm convergent in mean square and with probability one, i.e.,

$$\text{Prob} \left\{ \lim_{k \to \infty} p_k = p^o \right\} = 1. \tag{46}$$

The use of accelerated convergence methods can be easily formulated from existing control literature. The philosophy of using stochastic approximation was given in this section. The next section considers the use of the estimator for the autofocus problem.

V. Estimation in Autofocusing

An active autofocusing scheme has been described in reference [11]. Basically this scheme consists of adapting the focal length by sinusoidally perturbing the distance between the secondary mirror and the primary mirror of a Cassegrain telescope. This adjustment continues until the intensity is maximized. The scheme in [11] can be classified as a sinusoidal perturbation adaptive controller as previously derived in this paper. This section gives the results of application of the estimation to this problem.

The measurement equation can be functionally written as

$$y = k(z)\{I(\bar{d}) + \frac{\partial I}{\partial d} \sin \omega_d t + s_d + T + g\} + \eta \tag{47}$$

where d is the distance between the primary and secondary mirror (see Fig. 3). The models are as given previously for the speckle return, intensity, and gradient.

The simulations were conducted by using well known analytical expressions for the intensity of a focused, aperatured Gaussian laser. Thus, the intensity

at $\bar{d}$ was calculated by use of the analytical expression for I in reference [12] and the expression relating the focal length to $\bar{d}$. An analytical expression for $\partial I/\partial d$ may be obtained from

$$\frac{\partial I}{\partial \bar{d}} = \frac{\partial I}{\partial f}\frac{\partial f}{\partial \bar{d}} \tag{48}$$

where $\partial I/\partial f$ may be explicitly calculated from reference [12] and $\partial f/\partial d$ may be easily determined. The speckle noise was simulated using the shaping filter model as in equation (24) with appropriate white driving noise. These are used to form the measurements in the simulations. The estimator is then used to determine the best estimates for the gradient and speckle. The square roots of the diagonal elements of the covariance matrix gives the measure of the standard deviation of the estimation error. These are plotted in Figures 5-6. In addition, 200 Monte Carlo runs were conducted to ascertain the effects of the nonlinearities in the system dynamics. See Figure 4. Figure 7 yields the center value of the separation distance between the primary and secondary mirrors.

The modeling approach is thus based on Eq. (12), where k(z) is an effective range and target scattering dependent intensity scaling factor, $I(\bar{P})$, is based on the autofocus case discussed in Ref. [12], and speckle induced factors, s_d, T and g are derived for the target range and dynamics associated with that example. The inclusion of the speckle related interactive effects has not been considered earlier, and allows the autofocus example to be more realistic in terms of a practical implementation example.

The physical variables used in the simulations were

z_N = 1.1 KM
f_N = 800 M
r_s = 0.4 M (secondary radius of curvature)
r_p = 1.98 M (primary radius of curvature)
d_N = 0.7912 M (mirror spacing)
d = 0.02 M (dither length)
f_d = 1K Hz (dither focus rate)
$\omega_d = 2\pi fd = 2000\pi$
a = 0.0589 M
b = 0.299 M
λ = 10.6 μM

ℓ_T = target apparent brightness width
= 1 M
$\dot{\theta}_T$ = 10 mr/s → 100 mr/s
apparent angular rotation rate
β_g = 5 Hz
β_T = 30 Hz
β_d = 5 Hz
β_v = 5K Hz
$\sigma_g = \sigma_T$ = 10
σ_d = 100
σ_v = 0.1 .

The variance due to speckle (σ_d) at dither carriers is large due to changes in apparent rotation and apparent aspect angle changes with flight. The large magnitude for β_v indicates little temporal correlation from update to update for the correction loop. The estimator update rate was 4K Hz while the dither focus was driven at f_d = 1K Hz.

It is assumed that the estimator uses a range-doppler tracker for handover giving a value of range and a value of the rotation rate as

$$\dot{\theta} \cong \delta_{\text{doppler}} \lambda/\ell_{TN}$$

where

$$\ell_{TN} = f_N [2.44 \lambda/2b]$$

and f_N may be calculated from

$$d_N = \frac{r_s}{2}\left[\frac{1 - 2 f_N\left(\frac{1}{r_p} - \frac{1}{r_s}\right)}{2\left(\frac{f_N}{r_p}\right) - 1}\right] .$$

The problem of an uncertain rotation rate will not be treated here. It suffices to say that adaptive estimation techniques may be used to adapt upon this variable.

Returning now to the estimator equations, it can be noted that in order to estimate g, T and s, information must be available of the target apparent angular motion, $\dot{\theta}$, and brightness length, ℓ_T. The latter can be derived from the previous consideration. Focus estimation can be determined from the dither focus control servo (values of $f(d_o)$ which optimize $I(z,f)$. Angular rates can be derived from either a doppler measurement, or speckle induced frequency components at the sensor. The autofocus simulation assumed that initial handoff is from a range-doppler tracker, providing $z_N(o)$, $\dot{\theta}_N(o)$, $f_N(o)$ which are used to start the estimator. The complete description of the estimator then becomes

$$\begin{bmatrix} \dot{\hat{I}}(r) \\ \dot{\nabla_p \hat{I}} \\ \dot{\hat{s}} \\ \dot{\hat{x}}_2 \\ \dot{\hat{V}} \\ \dot{\hat{T}} \\ \dot{\hat{g}} \end{bmatrix} \begin{bmatrix} \varepsilon_k \nabla_p \hat{I}^T \nabla_p \hat{I} \\ -\hat{v} \\ \hat{x}_2 \\ -\beta_d \hat{x}_2 - \omega_d^2 \hat{s} \\ -\beta_V \hat{V} \\ -\beta_T \hat{T} \\ -\beta_g \hat{g} \end{bmatrix} + K \left\{ y - K(z)\left[1, \Omega^T, 1, 0, 0, 1, 1\right] \begin{bmatrix} \hat{I}(p) \\ \nabla_p \hat{I} \\ \hat{s} \\ \hat{x}_2 \\ \hat{V} \\ \hat{T} \\ \hat{g} \end{bmatrix} \right\}$$

The performance was compared with a simple (nearly ideal) phase lock loop.

VI. Conclusions

This paper generalizes the multidither adaptive optics concepts by showing that they belong to the class of sinusoidal perturbation adaptive control problems. This analysis yields insight into the higher order frequency terms as these can be easily seen to be related to the second derivatives of the intensity. This signal is corrupted by speckle effects as well as detector noise. The speckle components are modeled in state space formed by the use of the temporal power spectral density derived in this paper. The models are then used to form an extended Kalman estimator to estimate the

gradient vector. This separates the gradient from the speckle components.

The estimator may be used in a high bandwidth digital signal processor similar to digital radar processors in order to estimate the required quantities. It may be also used in an analog realization. This may lead to an elimination of conventional electronics to accomplish the necessary detection while eliminating the speckle noise from the measurement. This yields a control variable that is less noisy and, thus, better for feedback purposes. The control philosophy using stochastic approximation is briefly discussed. This yields an excellent convergent algorithm. The estimation algorithms were applied to the autofocus system. Stable convergence occurred in the presence of speckle induced competing signals. A similar case was tested with no estimation employed, leading, in general, to poor convergence for Monte Carlo simulated runs.

References

1. J. D'Azzo and C. Houpis, Feedback Control System Analysis and Synthesis, McGraw-Hill, New York, 1966.

2. A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

3. A. Gelb, Applied Optimal Estimation, M.I.T. Press, Cambridge, Mass., 1974.

4. J. Meditch, Stochastic Linear Estimation Control, McGraw-Hill, New York.

5. R. B. Asher and R. D. Neal, "Adaptive Estimation of Aberration Coefficients in Adaptive Optics," Proceedings of the 1976 IEEE Conference on Decision and Control, Submitted to the IEEE Trans. on Auto. Contr.

6. J. E. Pearson, et.al., "COAT Measurements and Analysis," RADC-TR-75-303, November 1975, RADC, OCTM GAFB, New York.

7. L. I. Goldfischer, J. Opt. Soc. Am. 55, 247 (1965).

8. R. B. Crane, J. Opt. Soc. Am. 60, 1658 (1970).

9. Ya Z. Tsypkin, Adaptation and Learning in Automatic Systems, Academic Press, New York, 1971.

10. R. B. Asher, D. Andrisani II, and P. Dorato, "Bibliography on Adaptive Control Systems," Proc. of the IEEE, pp. 1226-1240, August 1976.

11. A. Erteza, "A Method of Active Autofocusing Using an Apertured Gaussian Beam," Laser Digest, AFWL-TR-74-241, September 1974.

12. R. B. Asher, L. Sher, and R. D. Neal, "Focal Length and Target Range Estimation for Focused, Apertured Gaussian Laser Beams," to be submitted to Applied Optics.

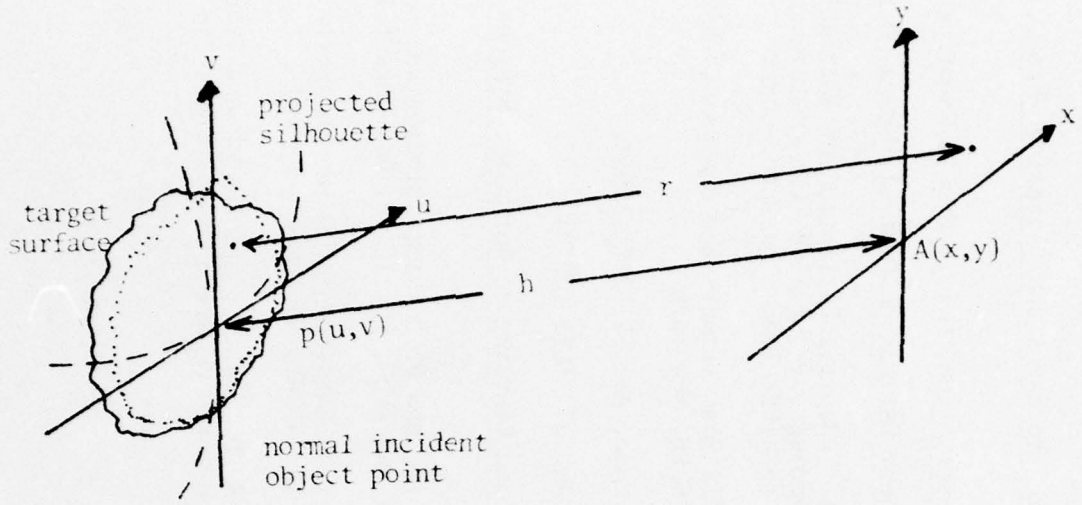


Fig. 1. Geometry of Illuminated Object.

Geometry of illuminated object with illumination function projected to the (u,v) plane, and the receiver plane (x,y) located h from the object.

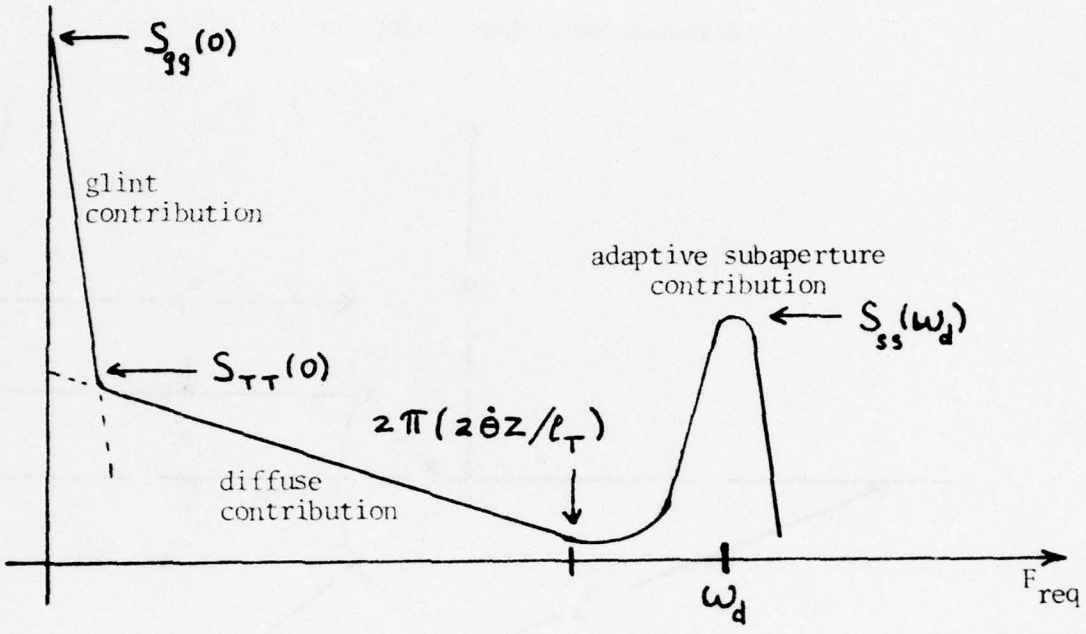


Fig. 2. Power Spectral Density.

Power spectral density taking into account target glint (specular regions), distributed diffuse regions and cross correlated adaptive subaperture factors. Only a point receiver is considered.

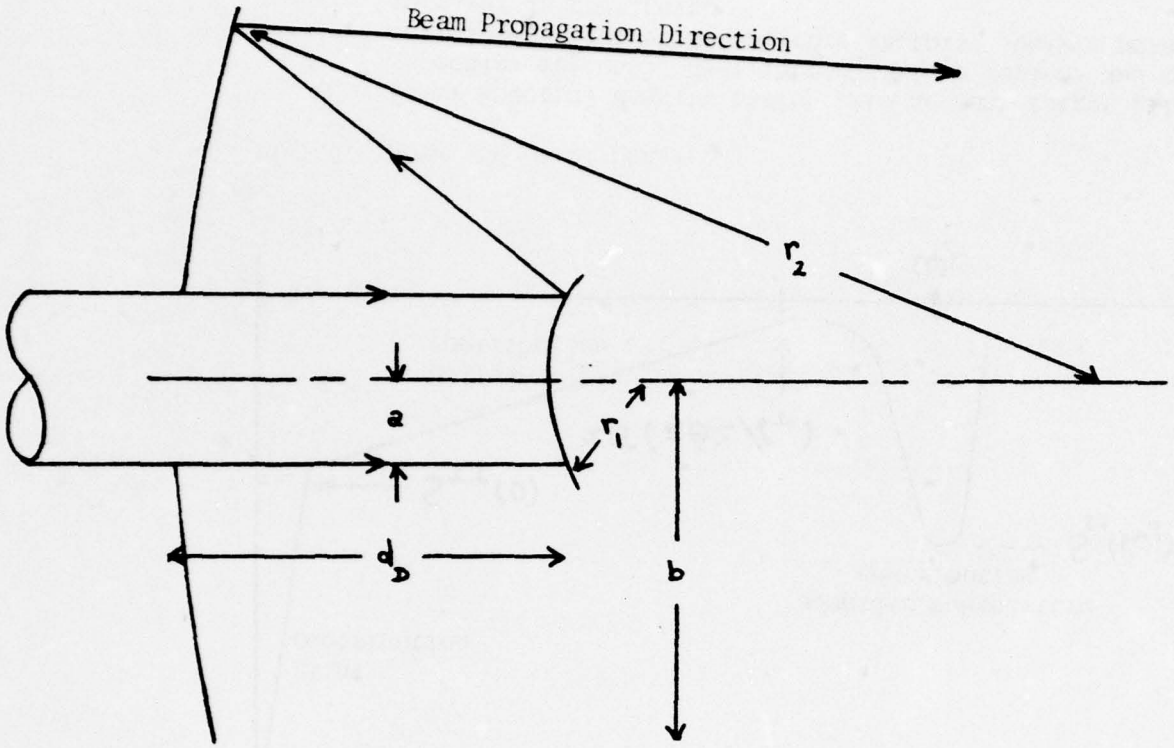


Fig. 3. Autofocus Cassegrain Geometry.

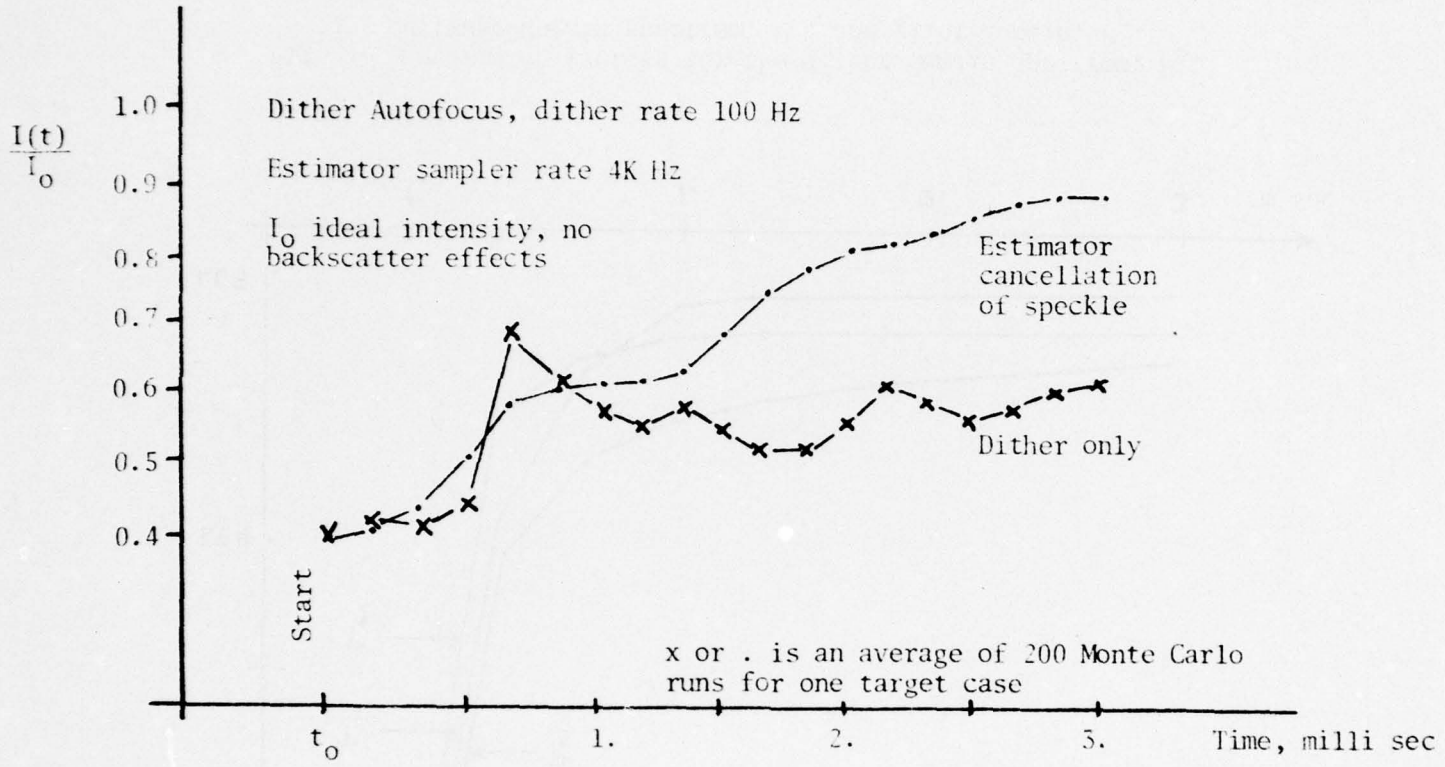


Fig. 4. Target Intensity - Autofocus Performance - Fixed Aspect.

Nose-on missile scenario. Multidither with and without estimation cancellation of speckle noise spectrum.

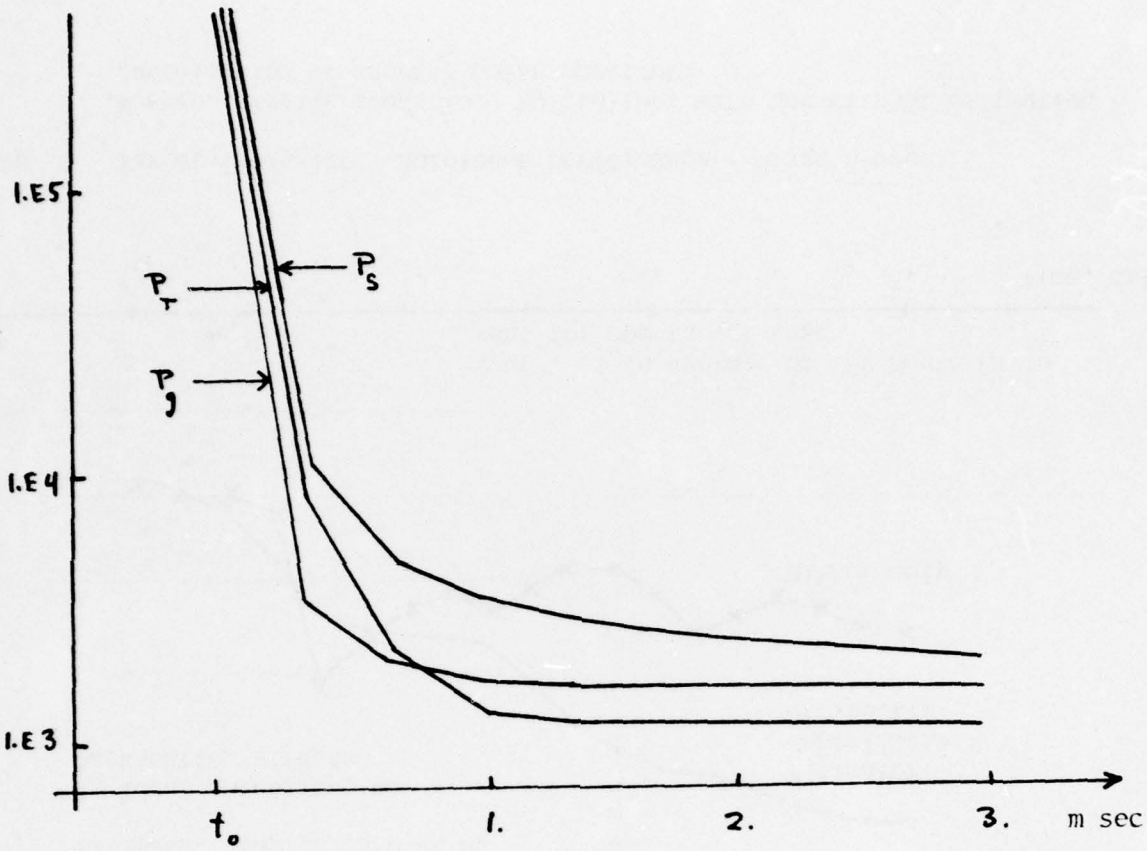


Fig. 5. Covariance factors for the dither domain spectrum, P_S; diffuse domain spectrum, P_T; and glint domain, P_g.

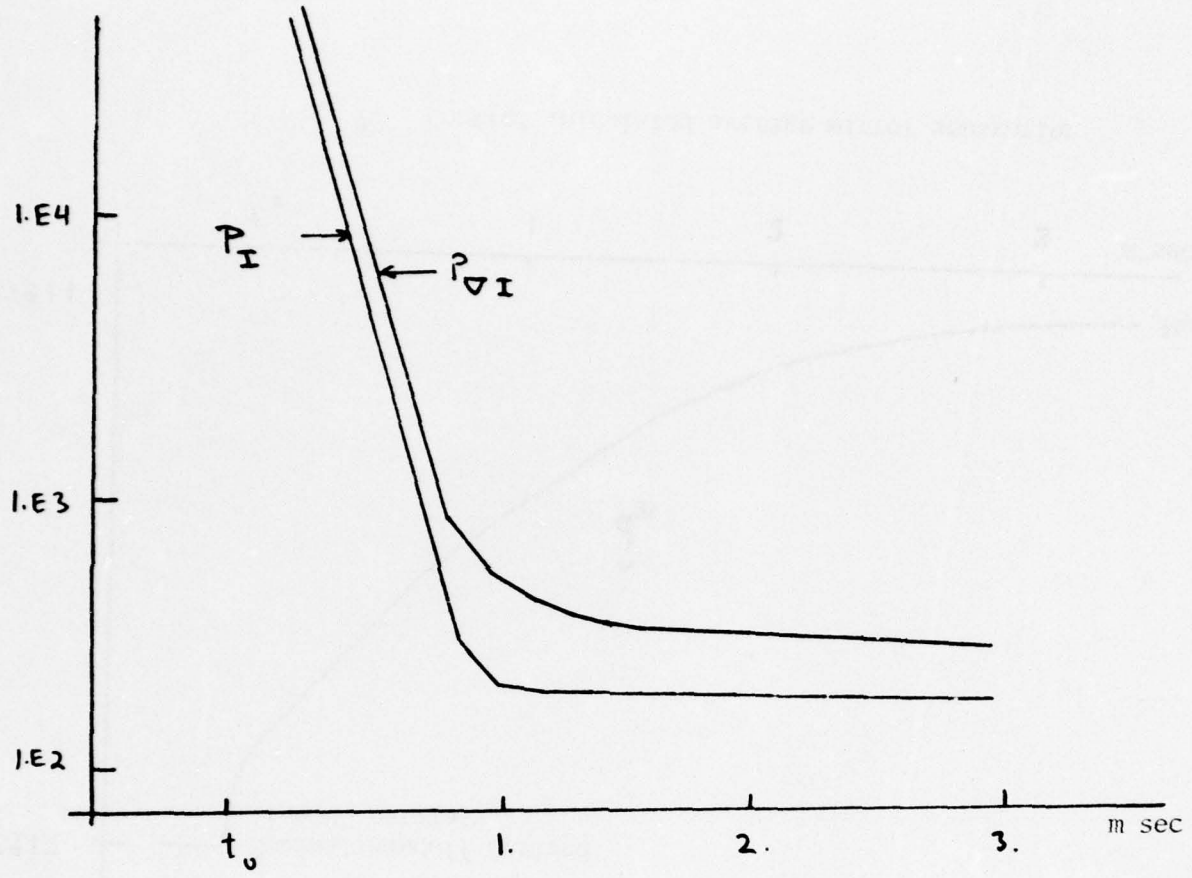


Fig. 6. Covariance for estimator intensity $\hat{I}$, and slope maximization, $\nabla\hat{I}$.

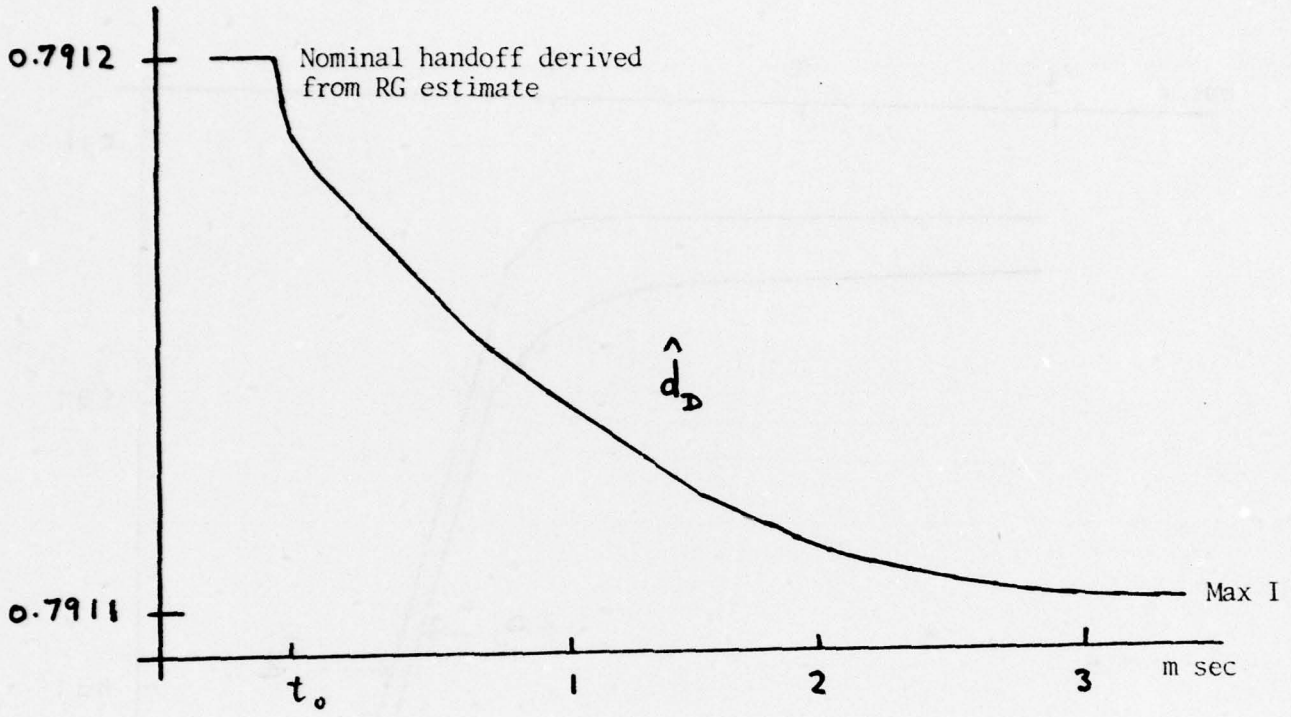


Fig. 7. Control calculated average mirror separation